Appendix B
Procedures For Deriving The Transition Matrix

This appendix explains the derivation of the cell counts for each column of the transition matrix. The notation here is as shown in Exhibit 6 of the report.

Variables Used from Food Stamp Quality Control Data

The model makes use of the following variables for each sample case in the food stamp QC database:

- STATUS: case error finding for review month
  - 1   amount correct
  - 2   overissuance ($25 or more)
  - 3   underissuance ($25 or more)
  - 4   ineligible

- ACTNTYPE: action type, for most recent action
  - 1   initial certification
  - 3   recertification
  - 5   interim change

- CERTMTH: length of assigned certification period, in months

- LASTCERT: number of months since last formal certification for food stamps (initial certification or recertification).

- RCNTACTN: date (year, month, and day) of the most recent action (either initial certification, recertification, or interim change)

- OCCDATE: date (year and month) of error occurrence (for error cases)[1]

- YRMONTH: date (year and month) for which the case has been sampled and reviewed (review month)

- FSEARN: monthly household earnings, as indicated in the case record (sum of wages and salaries, self-employment income, and other earned income).

It is important to note several limitations of the QC data, which require the creation of additional variables or the use of assumptions in estimating the model:

---

[1]   This variable is drawn from the "detailed error findings" recorded on the Quality Control Review Schedule, and is available only in the backup files made available for this analysis. The date of occurrence is recorded for up to nine errors that may have been discovered during the QC review.

- For cases in which the most recent action is an interim change, the variable LASTCERT indicates the number of months since the most recent *formal certification* (initial certification or recertification). Because of the need to establish the time elapsed since the most recent *action* (initial certification, interim change, or recertification) for all cases, an additional variable, LASTACT, is computed. This is the number of months since the most recent action, computed as the difference between RCNTACTN and YRMONTH. LASTACT thus differs from LASTCERT for those cases whose most recent action was an interim change.

- For active cases found in error at the QC review, the indicated timing of error occurrence is used to establish whether the case was correct or in error during the preceding month. An error case is considered to have been correct in the preceding month if the occurrence dates for all indicated errors are within two months of the review month. Conversely, an error case is considered to have been in error in the preceding month if the occurrence date for any of its indicated errors is more than two months before the review month.[2]

- For cases that are correct at the QC review, there is no indication in the QC record of the presence of error in any previous month. As indicated below, *we assume that an ongoing correct case was also correct in the preceding month, unless the case has just undergone an interim change*. The interim change is assumed to have occurred because the case was in error and required a benefit adjustment.[3]

The following sections of this appendix explain the derivation of the cell counts for each column of the matrix, beginning with the right-most of the five columns and then working to the left, using the notation in Exhibit 6 of this report. Using the definitions described below, and as shown in Exhibits B-1 through B-5, all cases in the QC sample are assigned to one (and only one) cell in columns two through five. Then, as described in the final section below, the cells in the first column entries are computed as row residuals.

---

[2] We also tested the model under alternative assumptions regarding this time window. It appears that reviewers may use the occurrence date to indicate the date of the underlying change in household circumstances that precipitated the error, rather than the date of onset of the QC error itself. Once a change in circumstances occurs, there is an allowable time interval for the client to report the change and a further time interval for the agency to act on the reported change.

[3] This assumption may overstate the number of cases previously in error, to the extent that: (a) an interim change may serve to prevent an error from arising, through timely benefit adjustment; or (b) an interim change may be recorded even though the caseworker simply processed a monthly or quarterly report without a change in benefit. However, the assumption may also understate the number of cases previously in error, by not taking account of "transient" errors that "self-correct"—i.e., errors associated with unreported short-term changes in household circumstances, such as temporary increases in income. Such cases "self-correct" (without an interim change) when the household's circumstances return to the *status quo ante*.

In this appendix, we refer to month-to-month transitions that occur between the QC review month and the prior month. In the body of the report, for ease of exposition, we adopt a different terminology, referring to transitions between the current month and the next month. We find it helpful to consider these transitions as occurring from March to April. Under this nomenclature, the review month is the next month, and the month preceding the review is the current month.

Expiring Error Cases

The fifth column of the transition matrix (shown in Exhibit 6) is comprised of error cases that are at the end of their certification period in the review month. By definition, such cases are already in error as they enter recertification. The column total (C_5) is the count of all cases that are in error (STATUS=2, 3, or 4) and whose certification is about to expire or has expired (LASTCERT $CERTMTH-1).[4] This includes, for example, error cases assigned a 12-month certification period and for whom 11 or more months have elapsed since their most recent certification.[5] Throughout this analysis, we use the word "expiring" to indicate that the case's certification period is about to end, not that the case is itself about to terminate. As noted above, these cases are ones about to undergo recertification; their error has preceded the recertification action.

We determine the prior-month status of each case in the fifth column, and thus determine its placement in one of the five cells in the fifth column, as follows (and as shown in Exhibit B-1):

---

[4]   We tested alternative specifications of this relationship. The one shown above was found to provide the most plausible results, based on information known about the length of assigned certification periods. If the mean length of certification period is n, the number of cases subject to recertification in any given month should be approximately 1/n. One alternative tested and rejected was to consider a case to be expiring if LASTCERT $CERTMTH. Under this definition the share of total cases classified as "expiring" each month was less than 0.100. This proportion was implausibly low, given that the mean certification length is between 9.5 and 10.0. The need to calibrate the model under alternative specifications arose from the fact that the QC review schedule records only the month and year (but not the day) of key events in a case's history.

[5]   The model will tend to understate somewhat the number of cases that are due or overdue for recertification, for the following reason. In entering the variable LASTCERT, QC reviewers are instructed that "For households that are participating in months for which they have not been certified, enter the number of months beyond the end of the household's certification period." For instance, if a case did not undergo its scheduled recertification at the 12th month and is now participating at the 15th month, LASTCERT should be coded as 3, not 15. It is apparent, however, that not all reviewers follow this instruction, as there are cases for which LASTCERT exceeds CERTMTH.

Exhibit B-1: Derivation of Fifth-Column Cell Counts: Expiring Error Cases

Cell Count	Prior-month ("March") status	Review-month ("April") status	Criteria (based on variables from QC review, as defined in text)
Q_{15}	Nonparticipating	Expiring error	equals 0, by assumption; see text
Q_{25}	Ongoing correct	Expiring error	certification expiring (LASTCERT\$CERTMTH-1) and error in review month (STATUS=2, 3, or 4) and correct in prior month (OCCDATE\$YRMONTH-2)
Q_{35}	Ongoing error	Expiring error	certification expiring (LASTCERT\$CERTMTH-1) and error in review month (STATUS=2, 3, or 4) and correct in prior month (OCCDATE<YRMONTH-2)
Q_{45}	Expiring correct	Expiring error	equals 0, by assumption; see text
Q_{55}	Expiring error	Expiring error	equals 0, by assumption; see text

- ***nonparticipating*** in prior month—The Q_{15} cell count is 0 by definition, as the model assumes that one-month certification periods are negligible, implying that a nonparticipant would never in the next month be an active case at the end of its certification period.

- ***ongoing correct*** in prior month—The Q_{25} cell consists of expiring cases for whom an error has just occurred, as indicated by an interval of two months or less between the error occurrence month and the review month (YRMONTH-OCCDATE # 2) for the earliest indicated error in the case.[6]

- ***ongoing error*** in prior month—The Q_{35} cell consists of expiring error cases that were previously in error, as indicated by an interval of more than two months between the error occurrence month and the review month (YRMONTH-OCCDATE >2) for the earliest indicated error in the case.

- ***expiring correct*** in prior month—The Q_{45} cell count is 0 by definition, as we assume that cases at the end of their certification period in the review month could not also have been expiring in the prior month.

- ***expiring error*** in prior month—For the same reason cited above, the Q_{55} cell count is also 0 by definition.

Expiring Correct Cases

The fourth column of the matrix consists of correct cases that in the review month are at the end of their certification period. The column total (C_4) is the count of all expiring correct cases: cases that are both correctly paid (STATUS=1) and whose certification is about to expire or has expired (LASTCERT $ CERTMTH-1).

We determine the prior-month status of each case, and thus its contribution to the cell counts in the fourth column, as follows (and as shown in Exhibit B-2):

- ***nonparticipating*** in prior month—The Q_{14} cell count is 0 by definition (as with Q_{15}), for we assume no one-month certification periods. A case thus cannot be nonparticipating in one month and an expiring active case in the next month.

---

[6] Here, once again, we tested and rejected alternative specifications, including YRMONTH-OCCDATE#1. The latter definition was found to assign an implausibly large number of cases to Q_{35} and Q_{55}, resulting in large negative entries for the Q_{31} and Q_{51}. The specification adopted here appears consistent with the QC rules that call upon reviewers to disregard errors (or "exclude variances") associated with changes in client circumstances that occur in the month immediately prior to the review month.

Exhibit B-2: Derivation of Fourth-Column Cell Counts: Expiring Correct Cases

Cell Count	Prior-month ("March") status	Review-month ("April") status	Criteria (based on variables from QC review, as defined in text)
Q_{14}	Nonparticipating	Expiring correct	equals 0, by assumption; see text
Q_{24}	Ongoing correct	Expiring correct	certification expiring (LASTCERT$CERTMTH-1) and correct in review month (STATUS=1)
Q_{34}	Ongoing error	Expiring correct	certification expiring (LASTCERT$CERTMTH-1) and correct in review month (STATUS=1) and interim action in review month (ACTNTYPE=5 and LASTACT=0)
Q_{44}	Expiring correct	Expiring correct	equals 0, by assumption; see text
Q_{54}	Expiring error	Expiring correct	equals 0, by assumption; see text

- *ongoing correct* in prior month—The Q_{24} cell consists of expiring correct cases which were also correct in the prior month, as indicated by the absence of any recent interim change (ACTNTYPE = 1 or 3, or ACTNTYPE = 5 and LASTACT > 0). Under the model's assumptions, the only circumstance in which an expiring correct case is regarded as having been in error in the prior month is if the case has just undergone an interim change, as accounted for below in cell Q_{34}.

- *ongoing error* in prior month—The Q_{34} cell consists of expiring correct cases for whom an error has just been corrected, as indicated by a recent interim change (ACTNTYPE = 5 and LASTACT = 0).

- *expiring correct* in prior month—The Q_{44} cell count is 0 by definition (as with Q_{45} earlier), assuming that cases reaching the end of their certification period in the review month could not also have been an expiring case in the preceding month.

- *expiring error* in prior month—For the same reason cited above, the Q_{54} cell count is also 0 by definition.

The assumption regarding cell Q_{34} deserves further explanation. In deriving this cell count, we assume that the only scenario in which an ongoing or expiring correct case would have been in error the previous month is the situation in which the case has just undergone an interim action. We assume no self-correction of errors—i.e., no situations in which a case in error one month becomes correct the following month, in the absence of agency action. There are some "transient error" situations, however, in which self-correction might occur in the absence of an interim change. Consider, for instance, a case that is overpaid because of unreported income from a particular source. If the income from that source later returns to zero, the case becomes correct. Conversely, however, there will be some instances in which an interim change has occurred, but where the case has remained correctly paid throughout. For instance, if the interim change were to adjust the benefit by less than $25 for an eligible case, there would have been no previous error, because of the $25 error tolerance. Also, if the interim change was in response to a client's timely report of a change in household circumstances, an error would have been prevented, but the case would not have actually been in error in the prior month.

In the text of this report, we have tested alternative specifications and examined the sensitivity of the model's findings to this assumption.

Ongoing Error Cases

The third column consists of error cases that in the review month are not at the end of their certification period. The column total (C_3) is the count of all cases that are in error in the

review month (STATUS=2, 3, or 4) and whose certification has not expired (LASTCERT < CERTMTH-1). This includes, for example, error cases assigned a 12-month certification period and for whom less than 11 months have passed since their last certification.

We determine the prior-month status of each case, and thus its placement in the cells of the third column, as follows (and as shown in Exhibit B-3):

- *nonparticipating* in prior month—The Q_{13} cell consists of households that have just been initially certified, as indicated by an initial certification that coincides with the review month (ACTNTYPE = 1 and LASTCERT=0).

- *ongoing correct* in prior month—The Q_{23} cell consists of ongoing cases whose error has just occurred, as indicated by an interval of two months or less between the error occurrence month and the review month (YRMONTH-OCCDATE # 2) for the earliest indicated error in the case.

- *ongoing error* in prior month—The Q_{33} cell consists of ongoing cases for whom the error has not just occurred, as indicated by an interval of more than two months between the error occurrence month and the review month (YRMONTH-OCCDATE >2) for the earliest indicated error in the case.

- *expiring correct* in prior month—The Q_{43} cell consists of cases that have just been recertified (ACTNTYPE = 3 and LASTCERT = 0) and whose error has just occurred, as indicated by an interval of two months or less between the error occurrence month and the review month (YRMONTH-OCCDATE # 2) for the earliest indicated error in the case.

- *expiring error* in prior month—The Q_{53} cell consists of cases that have just been recertified (ACTNTYPE = 3 and LASTCERT = 0) and whose error has not just occurred, as indicated by an interval of more than two months between the error occurrence month and the review month (YRMONTH-OCCDATE > 2) for the earliest indicated error in the case.

Exhibit B-3: Derivation of Third-Column Cell Counts: Ongoing Error Cases

Cell Count	Prior-month ("March") status	Review-month ("April") status	Criteria (based on variables from QC review, as defined in text)
Q_{13}	Nonparticipating	Ongoing error	initial certification in review month (ACTNTYPE=1 and LASTCERT=0) and error in review month (STATUS=2, 3, or 4)
Q_{23}	Ongoing correct	Ongoing error	certification not expiring (LASTCERT<CERTMTH-1) and error in review month (STATUS=2, 3, or 4) and correct in prior month (OCCDATE$YRMONTH-2)
Q_{33}	Ongoing error	Ongoing error	certification not expiring (LASTCERT<CERTMTH-1) and error in review month (STATUS=2, 3, or 4) and error in prior month (OCCDATE<YRMONTH-2)
Q_{43}	Expiring correct	Ongoing error	error in review month (STATUS=2, 3, or 4) recertification in review month (ACTNTYPE=3 and LASTCERT=0) correct in prior month (OCCDATE$YRMONTH-2)
Q_{53}	Expiring error	Ongoing error	error in review month (STATUS=2, 3, or 4) recertification in review month (ACTNTYPE=3 and LASTCERT=0) error in prior month (OCCDATE<YRMONTH-2)

Ongoing Correct Cases

The second column of the matrix consists of correct cases that are not at the end of their certification period in the review month. The column total (C_2) is the count of all cases that are correctly paid (STATUS=1) and whose certification has not expired (LASTCERT < CERTMTH-1).

We determine the prior-month status of these cases, and thus their contribution to the cell counts in the second column, as follows (and as shown in Exhibit B-4):

- *nonparticipating* in prior month—The Q_{12} cell consists of households that have just been initially certified, as indicated by an initial certification that coincides with the review month (ACTNTYPE = 1 and LASTCERT=0).

- *ongoing correct* in prior month—The Q_{22} cell consists of ongoing cases for which no error existed in the preceding month, as indicated by the absence of any recent interim change (i.e., ACTNTYPE = 1 or 3, or ACTNTYPE = 5 and LASTACT > 0). The rationale for this is the same as for the Q_{24} cell.

- *ongoing error* in prior month—The Q_{32} cell consists of ongoing cases for which an error has just been corrected, as indicated by a recent interim change (ACTNTYPE = 5 and LASTACT = 0). Here the logic is the same as for the Q_{34} cell.

- *expiring correct* in prior month—The Q_{42} cell count is computed as the row residual ($R_4 - Q_{41} - Q_{43}$).[7]

- *expiring error* in prior month—The Q_{52} cell count is computed under the assumption that, for a case entering recertification in error, the probability of becoming a correctly paid active case in the following month is the same as for an error case not undergoing recertification, which equals ($Q_{32} + Q_{34}$)/R_3.

The assumption used in computing the Q_{52} cell count warrants further explanation, as it figures importantly in simulations of the effects of more frequent recertification. Because the QC data do not provide information on the prior-month error status of cases that are correct in the review month, some assumption is necessary for either Q_{42} or Q_{52}. Once one of these values is estimated, the other (along with Q_{41} and Q_{51}) can be computed as either a row or column residual.

---

[7]   The calculation of row and column residuals must proceed on a specified order. The value Q_{42} is the last to be derived, preceded by Q_{52} and the first-column cells.

Exhibit B-4: Derivation of Second-Column Cell Counts: Ongoing Correct Cases

Cell Count	Prior-month ("March") status	Review-month ("April") status	Criteria (based on variables from QC review, as defined in text)
Q_{12}	Nonparticipating	Ongoing correct	initial certification in review month (ACTNTYPE=1 and LASTCERT=0) and correct in review month (STATUS=1)
Q_{22}	Ongoing correct	Ongoing correct	certification not expiring (LASTCERT<CERTMTH-1) and correct in review month (STATUS=1)
Q_{32}	Ongoing error	Ongoing correct	certification not expiring (LASTCERT<CERTMTH-1) and correct in review month (STATUS=1) interim action in review month (ACTNTYPE=5 and LASTACT=0)
Q_{42}	Expiring correct	Ongoing correct	derived as $R_4 - Q_{41} - Q_{43}$; see text
Q_{52}	Expiring error	Ongoing correct	derived as $((Q_{32}+Q_{34})/R_3)*R_5$; see text

Several alternate assumptions were tested. One was to assume that cases entering recertification as correctly paid (i.e., the expiring correct cases) should emerge upon recertification with an error rate equal to that of cases emerging from *initial* certification (later defined as the "first-month error rate"). We found, however, that this assumption, when used to compute the Q_{52} cell count, resulted in negative first-column entries in the national-level matrix, for several of the fiscal years under analysis. We did not adopt this assumption (nor several others that yielded similar results), as negative cell entries also appeared among many state-level matrices.

The assumption used in deriving Q_{52} minimized the extent of anomalous (negative) cell counts. This assumption might appear to inadequately capture the effectiveness of recertifications in removing errors from the active caseload. To the contrary, as will be seen later, this assumption yields rates of "error survival" that are one-third or one-fourth as large among *expiring* error cases (error cases subject to recertification) as among *ongoing* error cases (error cases not subject to recertification). Stated otherwise, a recertification will be shown to reduce very substantially (i.e., by more than one-half) the likelihood that an active error case remains so the following month.

Nonparticipating Households

The first column is comprised of households not participating in the Food Stamp Program in the review month. This includes the overwhelming majority of total U.S. households. The column total (C_1) is computed as the residual after subtracting the other four column totals from the total population ($Q - C_2 - C_3 - C_4 - C_5$). None of the cell entries in this column can be derived directly from the active case QC data, as nonparticipating households are not observed in active case reviews, by definition. Although there is a negative case action QC system for food stamp denials and terminations, the associated review process is not sufficient to establish whether a terminated case was correctly or incorrectly paid in its last month of participation.[8]

We determine the prior-month status of each case, and thus its contribution to the cell counts in the first column, as follows (and as shown in Exhibit B-5):

- nonparticipating in prior month—The Q_{11} cell consists of households that were nonparticipants in both the review month and the preceding month, computed as the first-row residual ($R_1 - Q_{12} - Q_{13}$).

---

[8]   See Mills (1990).

Exhibit B-5: Derivation of First-Column Cell Counts: Nonparticipating Households

Cell Count	Prior-month ("March") status	Review-month ("April") status	Criteria (derived as row or column residuals, as explained in text)
Q_{11}	Nonparticipating	Nonparticipating	derived as R_1-Q_{12}-Q_{13}; see text
Q_{21}	Ongoing correct	Nonparticipating	derived as R_2-Q_{22}-Q_{23}-Q_{24}-Q_{25}; see text
Q_{31}	Ongoing error	Nonparticipating	derived as R_3-Q_{32}-Q_{33}-Q_{34}-Q_{35}; see text
Q_{41}	Expiring correct	Nonparticipating	derived as C_1-Q_{11}-Q_{21}-Q_{31}-Q_{51}; see text
Q_{51}	Expiring error	Nonparticipating	derived as R_5-Q_{52}-Q_{53}; see text

- ***ongoing correct*** in prior month—The Q_{21} cell consists of households that have just left the caseload having previously been correctly paid in the midst of a certification period, computed as the second-row residual ($R_2 - Q_{22} - Q_{23} - Q_{24} - Q_{25}$).

- ***ongoing error*** in prior month—The Q_{31} cell consists of households that have just left the caseload having previously been incorrectly paid in the midst of a certification period, computed as the third-row residual ($R_3 - Q_{32} - Q_{33} - Q_{34} - Q_{35}$).

- ***expiring correct*** in prior month—The Q_{41} cell consists of households that have just left the caseload having previously been correctly paid at the end of a certification period, computed as the first-column residual ($C_1 - Q_{11} - Q_{21} - Q_{31} - Q_{51}$).[9]

- ***expiring error*** in prior month—The Q_{51} cell consists of households that have just left the caseload having previously been incorrectly paid at the end of a certification period, computed as the fifth-row residual ($R_5 - Q_{52} - Q_{53}$).

Because the first-column entries are computed as row residuals, based on cell counts in the other columns that are subject to sampling error, it is possible for the model to yield residuals that are negative in value. That is, for any given row, it is possible for the summed entries of the second through fifth cells to exceed the corresponding row total. Such first-column negative values are typically small, and we have imputed them as zero. This imputation (an upward change from a negative value) requires an offsetting downward adjustment elsewhere in the first column, to leave the first-column total unchanged. We make this downward adjustment in cell Q_{11}, to minimize the effect on the transition probabilities and error rate indicators that are derived from the cell counts. ***In the exhibits of this report, all imputed zero values are identified by asterisks to distinguish them from entries that are zero by definition or by estimate.***

---

[9]   It is necessary to first derive Q_{51} before computing Q_{41}.